CLAIMS

What is claimed is:

1. A vocoder apparatus, comprising:

(a) a superframe buffer for receiving multiple frames of voice data;

(b) a frame-based voice encoder analysis module for extracting parametric voice data from each frame within the superframe buffer;

(c) a superframe encoder for receiving parametric voice data for a series of frames within the superframe buffer from the analysis module, wherein parametric voice data received from the analysis module is selectively quantized to produce voice data which is encoded into an outgoing digital bit stream for transmission;

(d) a superframe decoder for receiving and decoding a digital bit stream encoded with superframe voice data into quantized frame-based parameters; and

(e) a frame-based decoder synthesizer for receiving the quantized parameters for each frame and decoding the quantized parameters into a synthesized voice output.

2. A voice compression apparatus, comprising:

(a) a superframe buffer for receiving multiple frames of voice data;

(b) a frame-based encoder analysis module for analyzing characteristics of voice data within frames contained in the superframe to produce an associated set of voice data parameters; and

(c) a superframe encoder for receiving voice data parameters from the analysis module for a group of frames contained within the superframe buffer, for reducing by analysis data for the group of frames and for quantizing and encoding said data into an outgoing digital

bit stream for transmission.

Sub B1

3. A voice compression apparatus as recited in claim 2, wherein the analysis module is capable of receiving voice data parameters is selected from the group of voice encoders consisting of linear predictive coders, mixed-excitation linear prediction coders, harmonic coders, and multiband excitation coders.

4. A voice compression apparatus as recited in claim 2, wherein said superframe encoder includes at least two parametric processing modules selected from the group of parametric processing modules consisting of pitch smoothers, bandpass voicing smoothers, linear predictive quantizers, jitter quantizers, and Fourier magnitude quantizers.

5. A voice compression apparatus as recited in claim 2, wherein said superframe encoder includes a vector quantizer wherein pitch values within a superframe are vector quantized with a distortion measure responsive to pitch errors.

6. A voice compression apparatus as recited in claim 2, wherein said superframe encoder includes a vector quantizer wherein pitch values within a superframe are vector quantized with a distortion measure responsive to pitch differentials as well as pitch errors.

Sub B2

7. A voice compression apparatus as recited in claim 2, wherein said superframe encoder includes a quantizer of linear prediction parameters, wherein quantization is performed with a codebook-based interpolation of linear prediction parameters that employ different

interpolation coefficients for each linear prediction parameter, and wherein said quantizer operates in closed loop mode to minimize overall error over a number of frames

8. A voice compression apparatus as recited in claim 7, wherein said quantizer is capable of performing a line spectral frequency (LSF) quantization using said codebook-based interpolation.

9. A voice compression apparatus as recited in claim 8, wherein said codebook is created by means of a training database operated on by a centroid-based training procedure.

10. A voice compression apparatus as recited in claim 2, wherein said superframe encoder includes a pitch smoother wherein calculations are based on an onset/offset classifier.

11. A voice compression apparatus as recited in claim 2, wherein said superframe encoder includes a pitch smoother wherein pitch trajectory is calculated using a plurality of voicing decisions.

12. A voice compression apparatus as recited in claim 11, wherein said pitch smoother classifies frames into onset and offset frames based on at least four waveform feature parameters selected from the group of waveform feature parameters consisting of energy. zero-crossing rate, peakiness, maximum correlation coefficient of input speech, maximum correlation coefficient of 500 Hz low pass filtered speech, energy of low pass filtered speech, and energy of high pass filtered speech.

13. A voice compression apparatus as recited in claim 2, wherein said superframe encoder includes a bandpass voicing smoother for mapping multiband voicing decisions for each frame into a single cutoff frequency for that frame, wherein said cutoff frequency takes on one value from a predetermined list of allowable values.

14. A voice compression apparatus as recited in claim 13, wherein said bandpass voicing smoother performs smoothing by modifying the cutoff frequency of a frame as a function of the cutoff frequencies of neighboring frames and the average frame energy.

15. A voice compression apparatus as recited in claim 2, further comprising means for compressing aperiodic flag bits for each frame in a superframe into a single bit per superframe, which bit is created based on the distribution of voiced and unvoiced frames within the superframe.

16. A voice compression apparatus as recited in claim 2, wherein said superframe encoder includes a plurality of quantizers for encoding parametric data into a set of bits, wherein at least one of said quantizers employs vector quantization to represent interpolation coefficients.

17. A voice compression apparatus as recited in claim 2, wherein a superframe is categorized into one of a plurality of coding states based on the combination of voiced and unvoiced frames within the superframe, and wherein each of said coding states is associated with a different bit allocation to be used with the superframe.

18. A voice compression apparatus, comprising:

(a) a superframe buffer for receiving multiple frames of voice data;

(b) a frame-based analysis module for determining a set of voice data parameters for said voice data; and

(c) a superframe encoder for receiving unquantized voice data parameters for groups of frames within a superframes, said superframe encoder comprising

(i) a pitch smoother for determining pitch and U/V decisions for each frame of the superframe and extracts parameters needed for frame classification into onset and offset frames,

(ii) a bandpass voicing smoother for determining bandpass voicing strengths for the frames within the superframe and determines cutoff frequencies for each frame, and

(iii) a parameter quantizer and encoder for quantizing and encoding voicing parameters received from said analysis module, said pitch smoother, and said bandpass voicing smoother into a set of bits and encoding said bits into an outgoing digital bit stream for transmission.

19. A voice decoder apparatus, comprising:

(a) a superframe decoder for receiving an incoming digital bit stream as a series of superframes and decoding and inverse quantizing said superframes into quantized frame-based voice parameters; and

(b) a frame-based decoder for receiving said quantized frame-based voice parameters

and combining said quantized frame-based voice parameters into a synthesized voice output signal.

20. A method of decoding a parametric voice encoded data stream into an audio voice signal comprising the steps of:

(a) buffering a received parametric voice data stream having a plurality of pitch periods and loading said buffered frame data into a buffer;

(b) constructing an estimated spectrum of excitation within each pitch period by breaking down the frequency spectrum into regions based on cutoff frequency, wherein said construction comprises the steps of:

(i) computing Fourier magnitude for each region, wherein the resultant computed Fourier magnitudes for at least one of said regions is then scaled by a gain factor computed for that region,

(ii) computing phase within each region, wherein the resultant phase for at least one of said regions has been modified by use of a weighted random phase, and

(iii) converting said Fourier magnitude and said phase within each region to a time domain representation by the computation of an inverse discrete Fourier transform; and

(c) generating an analog voice signal from said time domain representation.

21. A method as recited in claim 20, wherein said regions through which the frequency spectrum is broken down into comprise:

(a) a lower region wherein Fourier magnitudes directly determine the spectrum;

(b) a transition region wherein Fourier magnitudes are scaled down by a linearly decreasing weighting factor that drops from unity to a nonzero positive value dependent on the cutoff frequency of the current frame; and

(c) an upper region wherein Fourier magnitudes are scaled down by a weighting factor dependent on the cutoff frequency of the current frame.

22. An up-transcoder apparatus which receives a superframe encoded voice data stream and converts it to a frame-based encoded voice data stream, comprising:

(a) a superframe buffer for collecting superframe data and extracting bits representing superframe parameters;

(b) a decoder for inverse quantizing the bits for each set of superframe parameters into a set of quantized parameter values for each frame of the superframe; and

(c) a frame-based encoder for quantizing the voice parameters for each of the underlying frames, mapping said quantized voice parameters into frame-based data, and producing a frame-based voiced data stream.

23. A down-transcoder apparatus which receives an encoded frame-based voice data stream and converts it into a superframe-based encoded voice data stream, comprising:

(a) a superframe buffer for collecting a number of frames of parametric voice data and extracting bits representing frame-based voice parameters;

(b) a decoder for inverse quantizing the bits for each frame of parameter into quantized parameter values for each frame; and

(c) a superframe encoder for collecting said quantized frame-based parameters for the

group of frames within the superframe, producing a set of parametric voice data, and quantizing and encoding said parametric voice data into an outgoing digital bit stream.

24. A vocoder method for encoding digitized voice into parametric voice data, comprising the steps of:

(a) loading multiple frames of digitized voice into a superframe buffer;

(b) encoding digitized voice within each frame of the superframe buffer by parametric analysis to produce frame-based parametric voice data;

(c) classifying frames as onset frames and offset frames by calculating pitch and U/V parameters within each frame of the superframe;

(d) determining a cutoff frequency for each frame within the superframe by calculating a bandpass voicing strength parameter for the frames within the superframe buffer;

(e) collecting a set of superframe parameters from the parametric analysis, frame classification, and cutoff frequency determination steps for the group of frames within the superframe;

(f) quantizing the superframe parameters into discrete values represented by a reduced set of data bits that form quantized superframe parameter data; and

(g) encoding quantized superframe parameter data into a data stream of superframe-based parametric voice data that contains substantially equivalent voice information to the frame-based parametric voice data, yet at a lower bit per second rate of encoded voice.

25. A vocoder method for producing digitized voice from superframe-based parametric voice data, comprising the steps of:

(a) receiving superframe-based parametric voice data in a superframe buffer;

(b) decoding and inverse quantizing the voice data within the superframe buffer to recreate a set of frame-based voice parameter values; and

(c) decoding the frame-based voice parameters with a frame-based voice synthesizer which decodes the frame-based voice parameters to produce a digitized voice output.